# Noninvasive neurostimulation reveals a causal role for left superior temporal lobe in speech adaptation

Ja Young Choi[1,2], Tyler K. Perrachione[1]

[1]Department of Speech, Language & Hearing Sciences, Boston University,
[2]Program in Speech and Hearing Bioscience and Technology, Harvard University

## Background

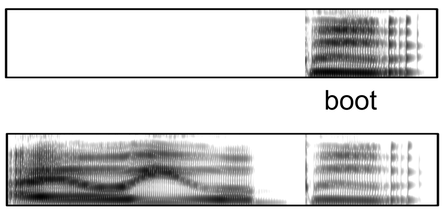
- Talker variability imposes additional processing cost, making listeners slower or less accurate when processing mixed-talker speech relative to single-talker speech [1-3].
- **Intrinsic talker normalization**: listeners use information contained within the speech sound to process the signal [4].
- **Extrinsic talker normalization**: extrinsic context can facilitate resolving the talker variability in acoustic-to-phonetic mapping [5].
- Neuroimaging studies have shown increased activity in STG when processing speech produced by mixed talkers relative to speech by one talker [6-8].
- In this study, we used **transcranial direct current stimulation (tDCS)** to investigate the **causal involvement of left STG in rapid talker adaptation using speech context** [9].

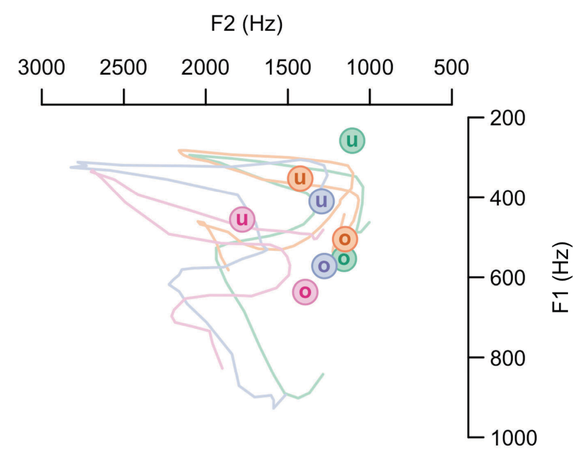
## Behavioral Task

**Participants:**
Native English speaking, right-handed adults with no history of speech, language, hearing or neurological disorder, or a significant head trauma (N=60; 46 female, 14 male; age 18-31, M=20.4 years)

**Stimuli:**
Recordings by 4 native English speakers (2 female, 2 male)

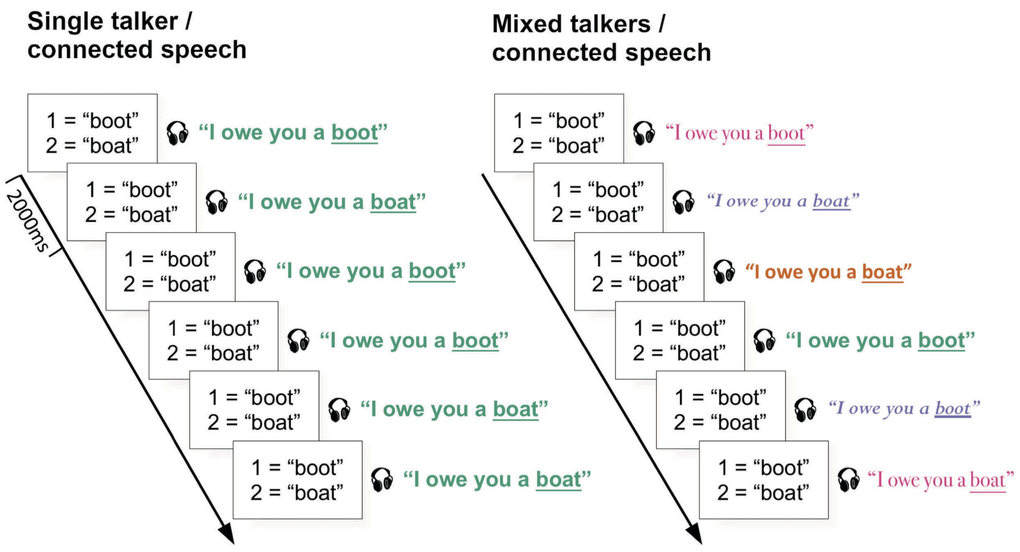
- **Target word:** "boot" / "boat"
- **Context:** "I owe you a …"



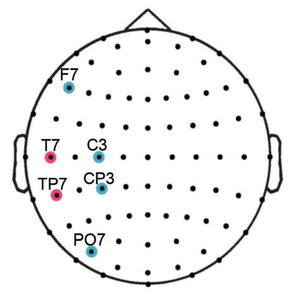

**Task design & procedure:**
Participants performed speeded word identification task that parametrically varied **talker variability** and **speech context**. They identified spoken words as quickly and accurately as possible.

- **Talker variability**: **single talker** vs. **mixed talkers**
- **Speech context**: **isolated words** vs. **connected speech**

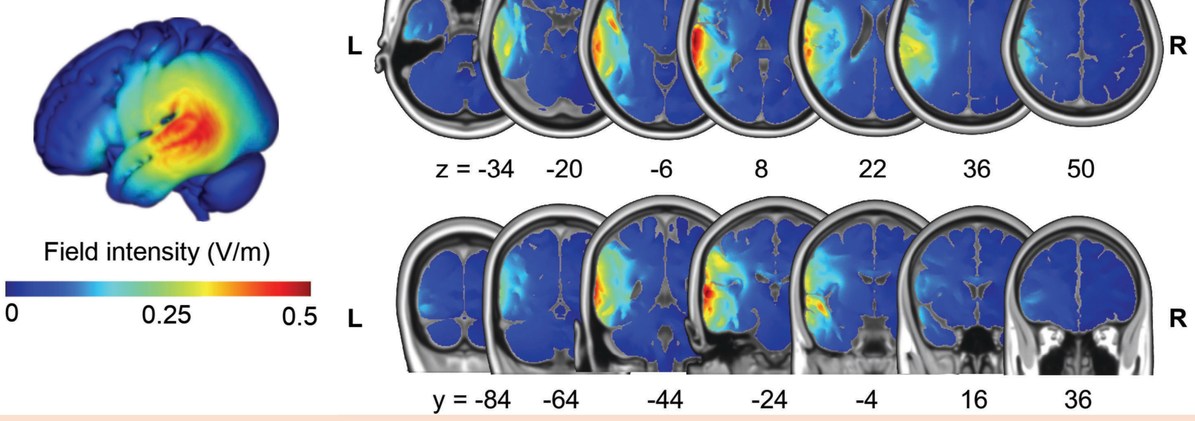



## Noninvasive Neurostimulation

- Participants were randomly assigned to anodal, cathodal, or sham HD-tDCS groups (n=20 in each group).
- Electrode configuration was chosen to target left STG.
  - **Stimulating** electrodes: T7, TP7
  - **Return** electrodes: F7, C3, CP3, PO7
- Participants received stimulation throughout the duration of the task except for the sham group.




### Simulated current flow maps




## Results

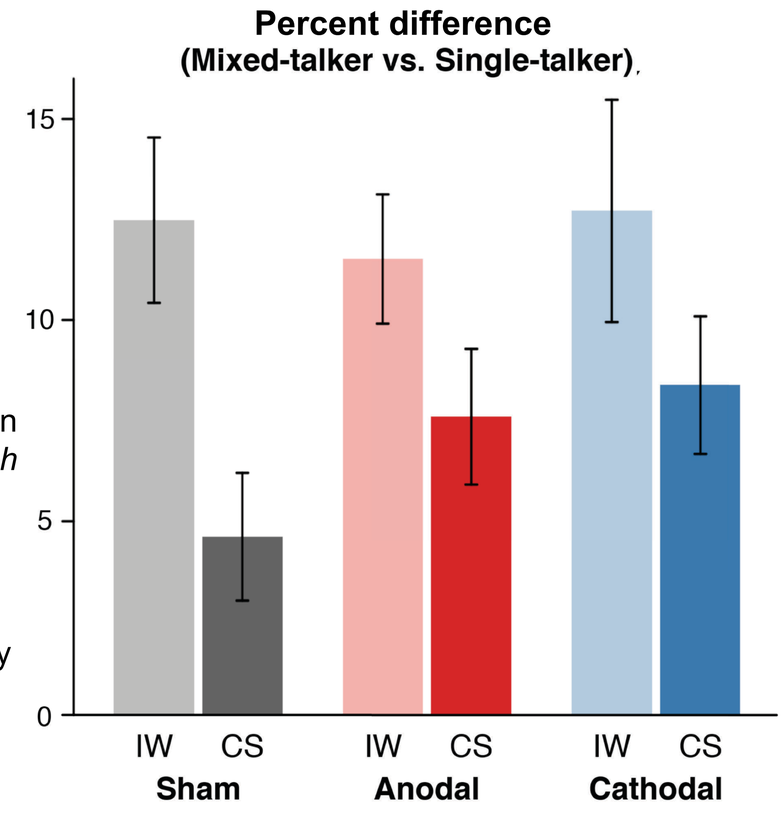
**Accuracy:** 98% ± 2%

**Effect of talker variability:**

- Response times in the mixed-talker condition were significantly slower than the single-talker condition ($F(1, 57) = 156.19$; $p \ll 0.001$).

**Effect of speech context:**

- Effect of talker variability was significantly smaller in the connected speech condition than in the isolated word condition (*speech context* × *talker variability* interaction; $F(1, 22275) = 89.74$; $p \ll 0.001$).
- Response times in the connected speech condition were significantly faster than the isolated word condition ($F(1, 57) = 98.15$; $p \ll 0.001$).




**Effect of stimulation:**

- The facilitatory effect of connected speech was reduced under both anodal and cathodal stimulation compared to sham condition (significant *stimulation* × *speech context* × *talker variability* interaction; $F(2, 22275) = 10.66$; $p < 0.01$)).

  Stimulation of left STG disrupted extrinsic talker normalization.

- Stimulation did not affect the effect of talker variability in the isolated word condition only (sham vs. anodal β = 0.014, SE = 0.018, t = 0.76, p = 0.45; sham vs. cathodal β = 0.0023, SE = 0.018, t = 0.13, p = 0.90).

  Stimulation of left STG did not have a significant influence on intrinsic talker normalization.

## Discussion

- The effect of extrinsic talker normalization was significantly reduced in anodal and cathodal stimulation relative to sham, revealing that **left STG is causally involved in extrinsic talker normalization**.
- However, stimulation did not affect intrinsic talker normalization. These results suggest a differential involvement of left STG in intrinsic and extrinsic talker normalization.
- Stimulation of left STG may be disrupting the precise balance between excitatory and inhibitory activity that enables rapid talker adaptation from preceding context.

## References

[1] Magnuson & Nusbaum (2007). *J. Exp. Psychol. Human*, 33, 391–409.
[2] Mullennix & Pisoni (1990). *Percept. Psychophys.*, 47, 379–390.
[3] Choi, Hu, & Perrachione (2017). *Attn. Percept. Psychophys.*, 80, 784-797.
[4] Nearey (1989). *J. Acoust. Soc. Am.*, 85, 2088-2113.
[5] Choi & Perrachione (2019). *Cognition,* 103982.
[6] Perrachione et al. (2016). *Neuron,* 92, 1383-1397.
[7] Wong, Nusbaum, & Small (2004). *J. Cog. Neuro.,* 16, 1173-1184.
[8] Zhang et al. (2016). *NeuroImage,* 123, 536-549.
[9] Choi & Perrachione (2019). *Brain Lang.,* 104655.

## Acknowledgments

We thank Elly Hu and Sara Dougherty. This research is supported by the NIDCD of the National Institutes of Health under award number R03DC014045 to TP.

Communication Neuroscience Research Laboratory
jayoung_choi@g.harvard.edu
http://sites.bu.edu/cnrlab/